Multipoint Control Protocol With Look-Ahead for Wavelength Division Multiplexed Ethernet Passive Optical Network

Xiaomin Liu, George N. Rouskas, Feng He, and Huagang Xiong

***Abstract*—We present a simple yet effective enhancement to the operation of the Ethernet passive optical network (EPON) multipoint control protocol (MPCP) for wavelength division multiplexing (WDM) EPON. The enhancement, inspired by our earlier work in a related but different context, allows the optical line terminal (OLT) to perform look-ahead scheduling on each of the upstream channels. The look-ahead operation is fully compatible with the existing standard and may be implemented via software updates to the OLT without affecting the operation of optical network units in EPON. The look-ahead enhanced MPCP achieves significant performance gains across a wide range of traffic loads and opens up new opportunities for the design of sophisticated dynamic bandwidth algorithms to support advanced quality of service capabilities.**

***Index Terms*—Look-ahead operation; Multipoint control protocol; Passive optical network.**

I. INTRODUCTION

The growth in network applications and services continues at a high rate, requiring network providers to offer improved and faster Internet connectivity so as to keep up with user demand. Technology constraints place a limit on the data rates that traditional access network architectures (e.g., DSL and cable) can support in connecting subscribers to the central office (CO). Therefore, fiber optic networks have an important role to play in the access network. Passive optical networks (PONs) are attractive due to their longevity, low operational costs, and high capacity and have already been deployed in the first/last mile [1].

In this work, we consider the Ethernet PON (EPON) and multiwavelength EPON technology that represents the dominant trend of PON deployment in the access network [2]. EPON is based on the point-to-multipoint architecture that is common to all PON technologies [3]. Specifically, EPON is deployed in a tree or tree-and-branch topology that connects an optical line terminal (OLT) to multiple optical network units (ONUs), typically via a $1{:}N$ splitter and $N{:}1$ combiner in the downstream and upstream directions, respectively. The OLT, located at the CO, is the root of the tree topology. The ONUs, residing at or near the customer premises, are the leaves of the tree topology and connect to user equipment. Communication between the OLT and the ONUs is carried out in a different mode depending on the direction of the transmission. In the downstream direction, that is, from the OLT to the ONUs, the EPON operates in a point-to-multipoint mode such that traffic from the OLT is broadcast to the ONUs. In the upstream direction, Ethernet packets from the ONUs to the OLT are time division multiplexed onto the single upstream wavelength. EPON is considered a shared media network in the upstream direction and uses the multipoint control protocol (MPCP) to manage and coordinate access to the shared upstream channel [3].

As the number of users increases and new bandwidth-intensive applications continue to emerge, it is expected that current single-channel EPONs will have to be upgraded in order to satisfy the growing traffic demands [1]. One approach to upgrading an existing EPON system that carries a high capital cost [4] is to increase the data rate from 1 to 10 Gbps. Another approach is to deploy multiple wavelengths, hence forming a wavelength division multiplexed (WDM) EPON. The ONUs may be equipped with fixed-tuned or tunable transceivers. This option allows for incremental upgrades, whereby the operator may add wavelengths as needed. A WDM EPON separates upstream data transmissions across both the time division multiplex (TDM) and wavelength dimensions but otherwise retains the overall EPON architecture.

In this paper, we present a simple yet effective enhancement to the operation of MPCP that results in a significant decrease, up to 70%, of packet delay across the whole range of traffic loads in WDM EPON systems. The enhancement, inspired by our earlier work in a related but different context [5], consists of a *look-ahead* operation that is managed by the OLT and allows the latter to coordinate access to the

---

Manuscript received March 5, 2013; revised October 15, 2013; accepted December 1, 2013; published January 10, 2014 (Doc. ID 186475).

Xiaomin Liu, Feng He, and Huagang Xiong are with the School of Electronics and Information Engineering, Beihang University, Beijing 100191, China.

George N. Rouskas (e-mail: rouskas@ncsu.edu) is with the Department of Computer Science, North Carolina State University, Raleigh, North Carolina 27695-8206, USA, and with King Abdulaziz University, Saudi Arabia.

http://dx.doi.org/10.1364/JOCN.6.000104

upstream channel in an efficient and effective manner. The look-ahead operation may be implemented via software updates to the OLT without affecting the operation of ONUs. With the proposed enhancement, the stable operation regime of a WDM EPON may be extended to high loads. The look-ahead feature also allows the OLT to dynamically allocate ONUs to upstream wavelength channels based on traffic demands. Such a load balancing operation adds to the effectiveness of the system in carrying time-varying demands.

The paper is organized as follows. In Section II, we discuss briefly the operation of MPCP at the media access control (MAC) layer of EPON, and we review related work on dynamic bandwidth allocation (DBA) schemes for both EPON and WDM EPON. In Section III we introduce the look-ahead enhancement to MPCP, and in Section IV we present simulation results to demonstrate the performance benefits that can be achieved. We conclude the paper in Section V.

II. MPCP AND RELATED WORK

MPCP, developed and standardized by the IEEE 802.3ah task force [6], is the protocol used to arbitrate the upstream transmission among the ONUs. MPCP does not dictate a specific DBA scheme, but it facilitates the implementation of DBA schemes by enabling the exchange of information that the OLT needs to allocate bandwidth to each ONU. MPCP introduces two 64-byte MAC control messages, `GATE` and `REPORT`. The OLT grants bandwidth to each ONU by sending a `GATE` message that informs the ONU of the start time and duration of its transmission on the upstream channel. Each ONU requests bandwidth by sending a `REPORT` message to the OLT that reports the current size of its transmission buffer. The two messages also carry timestamps that make it possible to determine the round-trip time (RTT) between the OLT and each ONU; the OLT uses the RTT information to ensure that the transmission windows of different ONUs do not overlap in time.

Several WDM EPON architectures have been proposed and studied; for a comprehensive survey, the reader is referred to [7]. In this work we consider the general (and interesting) case of a WDM EPON in which (1) the number of upstream channels is smaller than the number of ONUs and hence the ONUs have to share the upstream bandwidth and (2) ONUs have tunable transmitters. In this case, effective management of the upstream bandwidth requires the DBA to consider all channels in an integrated manner. In particular, there are two key problems that the DBA needs to address [8]: allocating ONU transmissions to the available upstream channels and arbitrating the transmissions on each channel. The latter issue is similar to that encountered in a single-channel EPON. To address the former problem, however, MPCP must be extended accordingly [1]. Specifically, we assume that the `GATE` message is modified to also include upstream channel information for the ONU. This information conveys to the ONU the upstream channel that it must use to transmit data during its next transmission window (i.e., the window specified by the start time and duration information included in the same `GATE` message); the ONU must then tune its transmitter to the specified channel before the start of its transmission window. This simple extension to MPCP enables the OLT to manage the upstream bandwidth effectively by scheduling upstream transmissions by ONUs on any wavelength channel. For instance, the OLT may track the utilization of the upstream channels and use the `GATE` messages to reallocate ONUs to wavelengths periodically to ensure that traffic is load-balanced across the various channels.

A. *Related Work*

Since all ONUs share the capacity of the common channel in the upstream direction, the development of efficient DBA algorithms that avoid collisions and attempt to optimize the utilization of the shared bandwidth resource has been a main focus of EPON-related research. For a comprehensive survey of the literature that reviews and classifies a wide range of DBA algorithms for EPONs, the reader is referred to [9]. In this section we only summarize the schemes that are most relevant to our work.

It was recognized early on [10] that bandwidth allocation schemes based on time division multiple access (TDMA) or basic polling would not be effective in an access network based on EPON technology: TDMA performs poorly under bursty IP traffic, while polling leads to high delays due to the accumulation of walk times. The interleaved polling with adaptive cycle time (IPACT) algorithm is an early scheme that improves upon basic polling, thus achieving high utilization [11,12]. According to IPACT, the OLT uses `GATE` messages to poll the ONUs in a round-robin fashion and grant each ONU a transmission window that reflects its backlog (as reported in the corresponding `REPORT` message). Two key ideas underlie the operation of IPACT. First, unlike basic polling schemes, the OLT does not poll each ONU sequentially; rather, it pipelines the `GATE` messages such that the walk times overlap and the idle time on the upstream channel is reduced significantly. Second, since the bandwidth grants reflect the instantaneous queue loads at the respective ONUs, the length of each transmission round adapts to the aggregate load on the upstream channel, and the bandwidth is allocated according to the requirements of each ONU, leading to effective statistical multiplexing.

The operation of MPCP with pipelining is shown in Fig. 1. As we can see, by pipelining the `GATE` messages, IPACT makes it possible to schedule the transmission windows of the various ONUs on the upstream channel so as to eliminate idle times (ignoring the guard bands between the transmission windows). The only idle time is between two consecutive rounds, as seen in Fig. 1 and as we discuss in more detail in the next section.

The basic IPACT algorithm has been extended and analyzed extensively in the literature. Variants of IPACT implementing different service disciplines can be classified as fixed service, gated service, or limited service [13–16].

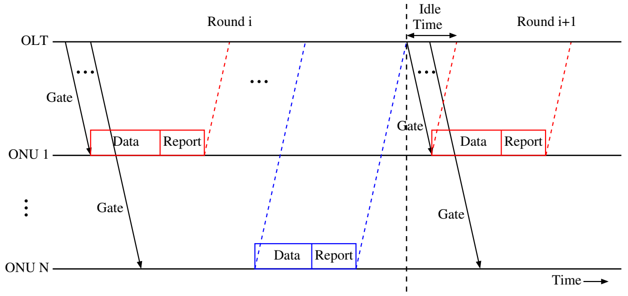

Fig. 1. Example of MPCP operation in EPON upstream bandwidth assignment.

Analytical models to compute the mean packet delay and mean queue length in an IPACT system using mean value analysis have been developed in [13,14]. DBA schemes that estimate the amount of new packets arriving between two consecutive polling instants and grant ONUs a larger window size based on this estimate were proposed in [15,16] and were shown to improve performance over the basic IPACT scheme. A polling protocol for EPONs called transmission upon reception was introduced in [17]; the protocol ensures collision-free transmission while taking into account fairness considerations in allocating bandwidth.

Polling mechanisms for WDM EPONs have also been developed and analyzed [18–20]. The WDM IPACT-ST (single polling table) algorithm in [18] is an extension of IPACT [11,12] that is designed for operation in a WDM environment. IPACT-ST assigns transmission windows to ONUs in a round robin manner such that they transmit in the first available upstream channel. Another polling mechanism called simultaneous and IPACT was presented in [19]. Furthermore, a group-synchronized polling algorithm was proposed in [20] for WDM EPONs.

DBA algorithms can generally be classified as online or offline [9]. With an online algorithm, as soon as the OLT receives a REPORT message from an ONU, it immediately schedules a transmission window for this ONU and sends the corresponding GATE message. With offline scheduling, the OLT waits until it receives REPORT messages from all ONUs; it then makes scheduling decisions and sends the GATE messages. An offline DBA uses traffic demand information from all the ONUs and can make scheduling decisions that are more effective in utilizing the upstream bandwidth than those of an online algorithm. On the other hand, waiting for all REPORT messages before making decisions may waste upstream channel resources. Online and offline scheduling algorithms also apply to WDM EPON. For instance, the next available supported channel online scheduling policy is adopted in [9]. An online just-in-time scheduling framework for WDM EPON was proposed in [21] to overcome the drawbacks of online and offline scheduling. The impact on polling of large differential distances between OLT and ONUs has been considered in [22], and various algorithms, including earliest finish time (EFT), latest finish time, EFT with void filling, and distance-based grouping were used to solve the grant scheduling problem. A generalized scheduling approach for both single-wavelength and WDM PONs, referred to as the K-out-of-N technique, was presented and analyzed in [23]; with this technique, efficiency is improved by having the OLT grant the top K requests out of the N ONUs for transmission within a round, where K is a tunable parameter. Finally, several DBA variants that use quality of service (QoS) and fairness criteria have been proposed in [24–28]; these schemes aim to support differentiated services and applications with heterogeneous requirements while making efficient use of the network resources. For an in-depth discussion of DBA algorithms for EPONs, please refer to [9].

The problem of assigning wavelengths to ONUs [7] has also been studied. In view of the potential for asymmetrical loads on the various channels, the study in [29] focuses on the fair assignment of excess bandwidth in the upstream direction. In considering network and service evolution, the work in [30] studies three types of WDM/TDM PONs, in terms of the number and tunability of ONU transmitters and introduces three respective lightweight optimal wavelength scheduling algorithms.

III. LOOK-AHEAD ENHANCED MPCP

A. Motivation

Our objective is not to introduce a new DBA scheme but rather to enhance the performance of the underlying MPCP protocol. The motivation for our work is based on the observation that it is possible to improve upon the basic MPCP pipelining scheme shown in Fig. 1. As the figure illustrates, there is an idle time on the upstream channel between two consecutive transmission rounds. This idle time is mandated by the fact that information carried by REPORT messages transmitted in round i is used to make bandwidth allocation decisions for round $i+1$. Specifically, when an offline DBA is employed, the OLT has to wait until it has received a REPORT message from each ONU in round i before it can finalize the bandwidth grants and send the first GRANT message in round $i+1$. For instance, the OLT has to ensure that the sum of bandwidth requests does not exceed a certain threshold on the length of a transmission round determined either by the specification or by desired bounds on, for instance, packet delay.

Let T_{idle} be the idle time between two consecutive rounds, $\text{RTT}_{\min}$ be the smallest RTT between the OLT and any of the ONUs, $T^{\text{OLT}}_{\text{proc}}$ be the time required by the OLT to process the REPORT messages and execute the DBA, and $T^{\text{ONU}}_{\text{proc}}$ denote the time required by the ONU to process the GATE message. Then, we have that

$$T_{\text{idle}} \geq \text{RTT}_{\min} + T^{\text{OLT}}_{\text{proc}} + T^{\text{ONU}}_{\text{proc}}. \tag{1}$$

The idle time T_{idle} increases packet delay and reduces the utilization of the upstream channel: if $\bar{R}$ is the average length of a transmission round, then channel utilization is $\bar{R}/(\bar{R}+T_{\text{idle}})$, not accounting for guard bands or other overhead that is independent of MPCP.

The motivation for our work is based on the observation that the idle time is due to the fact that existing approaches, including the one in [23], use requests made by ONUs in the *current* round for scheduling upstream transmissions. Next, we introduce a new look-ahead operation for MPCP, the novelty of which lies in the fact that bandwidth is allocated based on requests made in *earlier* rounds. This feature eliminates the idle time T_{idle} and, hence, is expected to improve the performance of any offline DBA scheme that is based on MPCP.

B. MPCP-ℓ: MPCP With Look-Ahead

We define MPCP-ℓ, an enhancement of MPCP that implements a look-ahead operation with parameter ℓ as follows:

Definition 3.1 (MPCP-ℓ): The MPCP protocol is configured such that queue length information carried by REPORT messages transmitted in round i is used by the OLT to allocate bandwidth in round $i + \ell$.

Clearly, when the look-ahead parameter $l = 1$, MPCP-1 is equivalent to the basic MPCP protocol.

Figure 2 illustrates the operation of MPCP-2, that is, when the look-ahead parameter $\ell = 2$. Queue length information carried in REPORT messages in round i is used by the OLT at the end of the round to execute a DBA method and allocate bandwidth for round $i + 2$. We make the reasonable assumption that DBA processing takes time less than the time for the ONUs to complete their transmissions in round $i + 1$; if that is not the case, we can increase the look-ahead value, as we discuss shortly. Therefore, during round $i + 1$, the OLT can start transmitting the GATE messages to inform ONUs of their transmission windows in round $i + 2$, as shown in Fig. 2. As long as the first such GATE message reaches the ONU before the end of round $i + 1$, the idle time between rounds is eliminated, ensuring continuous transmission on the upstream channel (ignoring, of course, guard bands or other gaps between frames mandated by Ethernet). In essence, the look-ahead operation masks the three components of the idle time in the right-hand side of expression (1), namely, minimum RTT between the first GATE message and receipt of the first bit of data, processing time for DBA, and processing time at the ONU, by overlapping them with data transmission during a round. As a result, this small change in the operation of MPCP completely removes the delays associated with the DBA and GATE messages.

This operation can be readily generalized to look-ahead values $\ell > 2$. Larger look-ahead values would be needed if the idle time from expression (1) is larger than the average length of a transmission round, that is, due to long RTTs between the OLT and ONUs and/or the processing requirements of the DBA. In this case, it may be necessary to use more than one round to completely mask the idle time.

The look-ahead feature of MPCP-ℓ, $\ell \geq 2$, improves the delay and throughput performance of the protocol by achieving better utilization of the upstream channel, as the numerical results we present in the next section indicate. The look-ahead operation affords further benefits. First, since the delay introduced by the RTT is masked, the allowable distance between the OLT and ONUs is limited only by transmission impairments, not performance issues due to the MPCP protocol. Second, whereas the execution time of DBA increases the idle time of MPCP, it does not affect the channel utilization of MPCP-ℓ, $\ell \geq 2$. Consequently, network operators may implement sophisticated bandwidth allocation algorithms that would not otherwise be possible to implement in basic MPCP due to running time constraints.

C. MPCP-ℓ for WDM EPON

The look-ahead feature of MPCP may also be implemented in a WDM EPON. We consider two cases. If all ONUs have a fixed-tuned transmitter, then each upstream channel operates independently, as we discussed earlier. Therefore, the OLT simply executes a separate instance of MPCP-ℓ for each upstream channel. In fact, due to channel independence, the parameter ℓ of the protocol may be different for different channels.

Consider now a WDM EPON in which each ONU has a tunable transmitter. In this case, the OLT executes a single instance of MPCP-ℓ. Specifically, upon receiving the REPORT messages from all ONUs within transmission round i, the OLT carries out three operations. First, it runs a load balancing algorithm to allocate wavelengths to

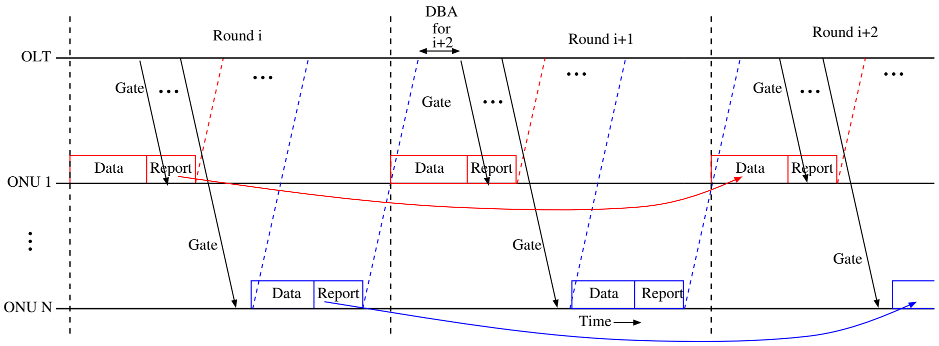

Fig. 2. Operation of MPCP-2 with look-ahead parameter $\ell = 2$.

ONUs, in other words, to determine the set of ONUs that will transmit on each upstream channel during round $i + \ell$. Then it runs a DBA algorithm for each channel to schedule the round $i + \ell$ transmissions by ONUs assigned to that channel. Finally, it transmits the scheduling information, including the upstream channel to be used, in the `GATE` message it sends to each ONU.

Note that it is the upstream channels (wavelengths) that are allocated to the various ONUs, not the downstream channel. Since the downstream wavelength does not change, the OLT may send `GATE` messages on this channel regardless of the upstream channel assignment; all ONUs receive the `GATE` messages correctly since they have a receiver fixed to the downstream channel. Once an ONU receives a `GATE` message for round i, and assuming that it must tune its transmitter between rounds $i - 1$ and i, it will do so immediately after transmitting its packets in round $i - 1$. Given that the length of a round is relatively large (e.g., 2 ms) compared to the transmitter tuning time, the ONU will have sufficient time to retune its transmitter between rounds when it is necessary to do so.

Note that, similar to the single-channel EPON, MPCP-ℓ prescribes neither a DBA algorithm nor a load balancing algorithm. For instance, the load balancing algorithm may attempt to spread the traffic demands in round $i + \ell$ evenly across all channels without any other considerations. Alternatively, the load balancing algorithm may take other parameters into account, for instance, balancing the traffic while also maximizing the number of ONUs that do not need to tune their transmitter from the channel they were assigned in round $i + \ell - 1$. The important point is that the look-ahead operation supports the development of more sophisticated algorithms than would be possible without it.

D. Look-Ahead Implementation Considerations

Implementing the look-ahead feature does not require any changes to the `GATE` messages. For `REPORT` messages, two relatively minor modifications are required at the OLT only, without any change to how ONUs operate. Specifically, the OLT uses the information in the `REPORT` messages to allocate bandwidth not in the next transmission round, as with basic MPCP, but in a future round determined by the value of the look-ahead parameter ℓ. Second, the OLT must be careful in how to interpret the queue length information reported by the ONUs since the latter record instantaneous queue lengths in their report. For instance, consider $\ell = 2$ as shown in Fig. 2. When, say, ONU 1 reports its queue length in round i, it includes all the packets currently in its queue. However, the OLT has already allocated bandwidth to ONU 1 in a `GATE` message that has not been received or processed by the ONU. Therefore, the OLT must subtract this allocated bandwidth from the queue length information carried in the `REPORT` message from ONU 1 before it uses it to allocate bandwidth for round $i + 2$. Finally, to jumpstart the look-ahead operation, for any value of ℓ, the OLT must initially transmit ℓ rounds of `GATE` messages that make bandwidth grants only large enough for the ONUs to transmit `REPORT` messages; from then on, the bandwidth grants will be determined from queue length information from ℓ rounds in the past.

IV. NUMERICAL RESULTS

A. Simulation Model

1) Network Environment: For the simulation results we present in this section, we consider a WDM EPON with one OLT and N ONUs, $N = 32, 64, 128$. The network has W upstream channels, $W = 1, \ldots, 16$; when $W = 1$, the system reduces to a single-channel EPON. Since we are only interested in the performance of traffic in the upstream direction, without loss of generality, we assume that there is only one downstream channel. Each ONU employs one fixed-tuned receiver (listening on the downstream channel) and one tunable transmitter capable of tuning to any of the upstream channels. We also assume that tuning times are negligible compared to the RTT and transmission windows. All channels (upstream and downstream) operate at a rate of 1 Gbps, and each channel is on a different wavelength. The distance between the OLT and the ONUs is within the range of 2–5 km. The `GATE` and `REPORT` messages each have a length of 64 bytes. The length of each transmission round can be no larger than 2 ms, and we assume a 5 μs guard time between successive transmission windows from different ONUs. The buffer size of each ONU is limited to 10 KB.

2) Traffic and Packet Length Distributions: Packet lengths L (in bytes) at each ONU are generated from a trimodal distribution that is meant to reflect the distribution of packet lengths that has been observed in the Internet [31]:

$$P[L = x] = \begin{cases} 0.4, & x = 40 \\ 0.2, & 41 \le x \le 1449 \\ 0.4, & x = 1500 \end{cases}. \tag{2}$$

Consequently, the average packet length is 770 B. We consider two types of traffic distribution, uniform and hot-spot. With uniform distribution, each ONU generates the same amount of traffic, whereas under the hot-spot distribution 25% (respectively, 75%) of the ONUs generate 80% (respectively, 20%) of the traffic.

3) Bandwidth Allocation: Recall that the focus of our work is on the look-ahead enhancement to MPCP, not bandwidth allocation. Therefore, in our simulations we use a simple strategy that allocates each ONU a bandwidth grant sufficient to satisfy the corresponding request, as long as the length of the transmission round is no larger than 2 ms. If the sum of bandwidth requests for a given round exceeds 2 ms, then the OLT scales down all the requests by a constant factor so that their sum does not exceed 2 ms and grants the corresponding amount to each ONU.

4) Load Balancing: The look-ahead operation is an instance of offline scheduling as defined in Subsection II.A, in that the OLT schedules transmissions in round i based on queue information submitted by the ONUs in round $i - \ell$. Since the OLT has complete information regarding the traffic demands for each transmission round i, it can assign transmission wavelengths to ONUs so as to ensure that the traffic is balanced across the upstream channels. To this end, we apply the largest processing time (LPT) algorithm [32], an approximation algorithm for the multiprocessor scheduling problem that works well in practice. In the EPON context, each channel corresponds to a processor, and the queue length of each ONU corresponds to a task that must be scheduled (transmitted) on one of the processors (channels). The objective of LPT is to minimize the finish time (makespan) of the schedule, and this objective is equivalent to load balancing across the processors.[1] Once the ONU has determined the channel assignment and start time for each ONU, it conveys this information in the modified GATE messages, as we described earlier.

B. Results: Single-Channel EPON

Let us first consider a single-channel EPON (i.e., $W = 1$). Figures 3–5 present the results of OPNET simulations comparing the performance of three protocols, MPCP-ℓ, $\ell = 1, 2, 3$, where MPCP-1 is equivalent to the original MPCP protocol and the other two implement two versions of look-ahead scheduling. These figures present results for a 32-node EPON; results for a 16-node network are very similar and can be found in [35]. For all the results, we have estimated 95% confidence intervals using the method of batch means. Since the confidence intervals are narrow, we omit them to avoid cluttering the figures.

Figures 3 and 4 plot the average packet delay as a function of the traffic load under uniform and hot-spot traffic, respectively. The main observation from the figures is that the look-ahead operation reduces the delay considerably compared to the original MPCP across the whole range of traffic loads. Despite the fact that all packets are kept in the queue for an additional amount of time equal to one or two transmission rounds under MPCP-2 and MPCP-3, respectively, look-ahead scheduling eliminates the idle time between successive rounds, resulting in lower delay per round, such that in a steady state the average packet delay is significantly lower.

We also note that the delay is higher for MPCP-3 than for MPCP-2. This result is due to the fact that, for the system parameters we used in the simulation scenarios, the idle time is smaller than the length of a transmission round; hence the extra delay that packets incur under MPCP-3 does not offer any extra benefit. As we mentioned earlier, a look-ahead parameter $\ell > 2$ would be of value when two

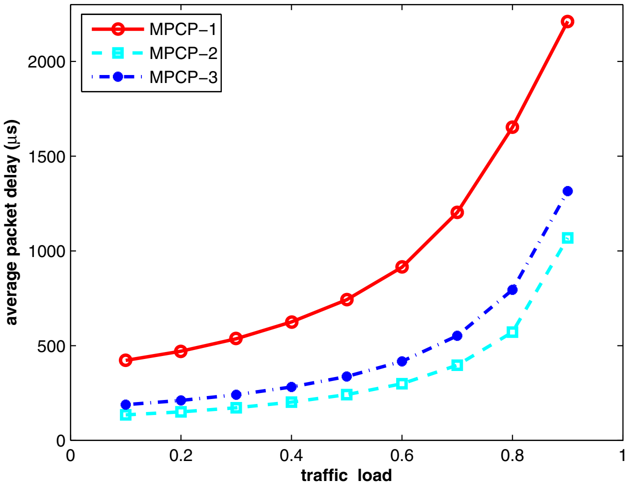

Fig. 3. Packet delay versus traffic load, with $N = 32$ and uniform traffic.

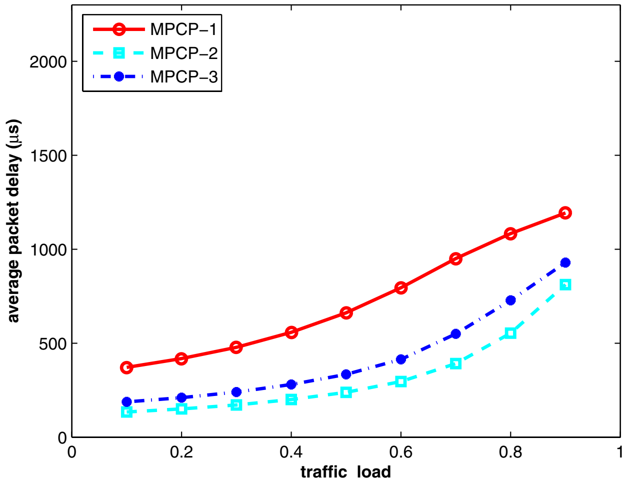

Fig. 4. Packet delay versus traffic load with $N = 32$ and hot-spot traffic.

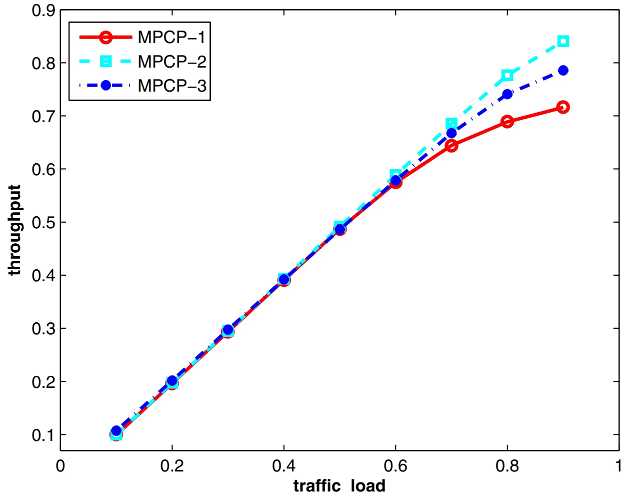

Fig. 5. Throughput versus traffic load, with $N = 32$ and hot-spot traffic.

---

[1] More sophisticated objectives may be considered in assigning upstream channels to ONUs, for instance, those similar to the ones we have in a similar context in [33,34]. However, in this paper we are simply interested in demonstrating the effectiveness of MPCP-ℓ, not in exploring new channel assignment schemes.

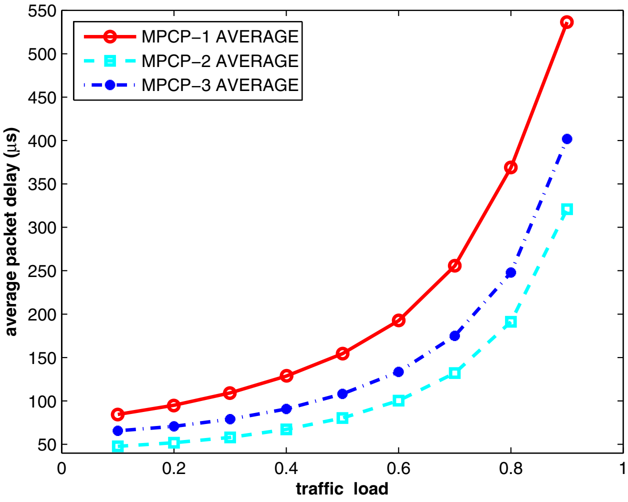

Fig. 6. Packet delay versus traffic load, $N = 32$, $W = 3$, uniform traffic.

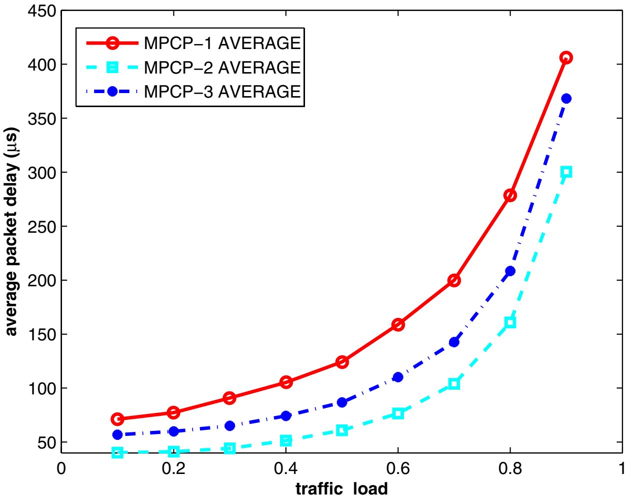

Fig. 8. Packet delay versus traffic load, $N = 128$, $W = 16$, uniform traffic.

or more transmission rounds are necessary to completely mask the idle time, that is, for networks with long RTTs or when the DBA is computationally expensive.

Another interesting observation from these figures is that delays under the hot-spot traffic scenario are lower than those under the uniform traffic scenario. This behavior can be explained by the fact that in the simulations we have the OLT schedule ONUs with large demands early in the transmission round. Since four out of sixteen ONUs generate 80% of the total traffic under the hot-spot scenario, scheduling this traffic early reduces the overall average delay.

Finally, Fig. 5 plots the aggregate throughput on the upstream channel as a function of traffic load. As we can see, the throughput increases almost linearly until the load reaches 60%. After that point, the overhead due to idle time on the operation of the original MPCP is evident in the fact that the corresponding curve increases more slowly. On the other hand, the look-ahead feature also improves the throughput performance of the protocol, and, especially for $\ell = 2$, the throughput curve increases almost linearly until the load reaches 80%.

C. Results: WDM EPON

1) Delay Under Uniform Traffic: Figures 6–8 plot the average packet delay as a function of traffic load under uniform traffic and three protocols, MPCP-ℓ, $\ell = 1, 2, 3$. Figure 6 shows results for a 32-node EPON with $W = 3$ wavelengths, Fig. 7 shows results for a 64-node EPON with $W = 8$, and Fig. 8 shows results for a 128-node EPON with $W = 16$ upstream channels; results for other combinations of the number N of nodes and number W of channels are similar and are omitted. For each protocol (i.e., value of l), the three figures plot curves for the overall average delay across the stated number of channels; curves for each

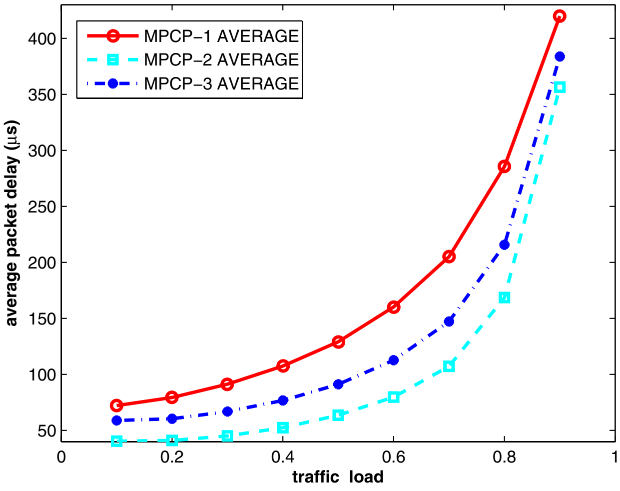

Fig. 7. Packet delay versus traffic load, $N = 64$, $W = 8$, uniform traffic.

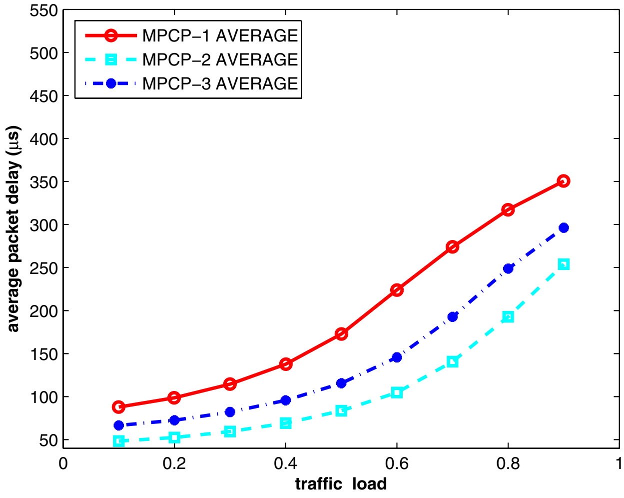

Fig. 9. Packet delay versus traffic load, $N = 32$, $W = 3$, hot-spot traffic.

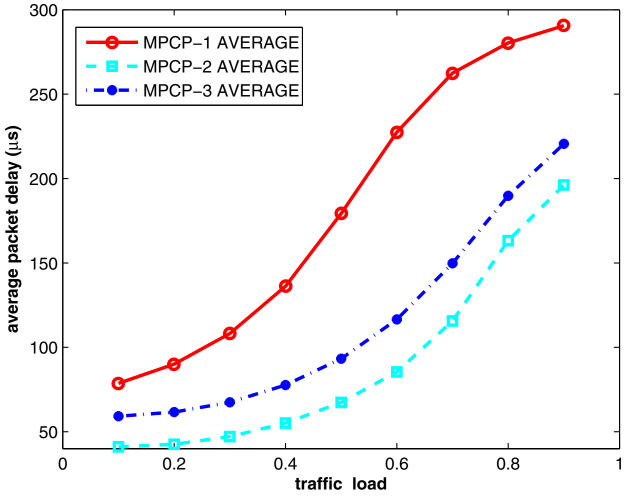

Fig. 10. Packet delay versus traffic load, $N = 64$, $W = 8$, hot-spot traffic.

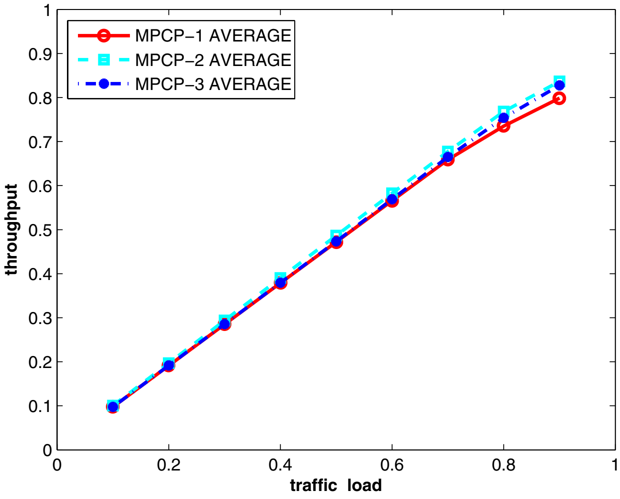

Fig. 12. Throughput versus traffic load, $N = 32$, $W = 3$, uniform traffic.

individual channel exhibit similar behavior and are omitted to avoid cluttering the figures.

In terms of the relative performance of the MPCP-ℓ protocols, the results are consistent with our earlier observations regarding the single-channel EPON: MPCP-2 results in the lowest overall average delay and MPCP-1 (i.e., the original protocol without look-ahead) in the highest. MPCP-3 performs better than MPCP-1 but worse than MPCP-2 for the same reasons we explained in the single-channel EPON above. These results indicate that the look-ahead operation is also effective in a multichannel network and that, under the network parameters considered in our study, MPCP-2 may reduce average delay up to 60%–70% across a wide range of loads compared to MPCP-1 that does not have the benefit of look-ahead.

2) Delay Under Hot-Spot Traffic: The three Figs. 9–11 are similar to the three Figs. 6–8, respectively, but present results for hot-spot traffic. Our observations regarding the relative performance of the three MPCP variants are identical to those regarding uniform traffic. Also, the average delays under hot-spot traffic are lower than those for uniform traffic at the same load, consistent with the results for the single-channel EPON case we presented earlier.

3) Throughput: Figures 12–14 plot the throughput on the upstream channels of a WDM EPON as a function of traffic load under uniform or hot-spot traffic and three combinations of the number N of nodes and number W of wavelengths; other such combinations result in similar behavior and are omitted. The figures plot the average throughput over all upstream channels of a given WDM EPON. Similar to our earlier observations regarding the single-channel EPON, MPCP-2 exhibits the highest throughput, followed by MPCP-3 and MPCP-1, in this

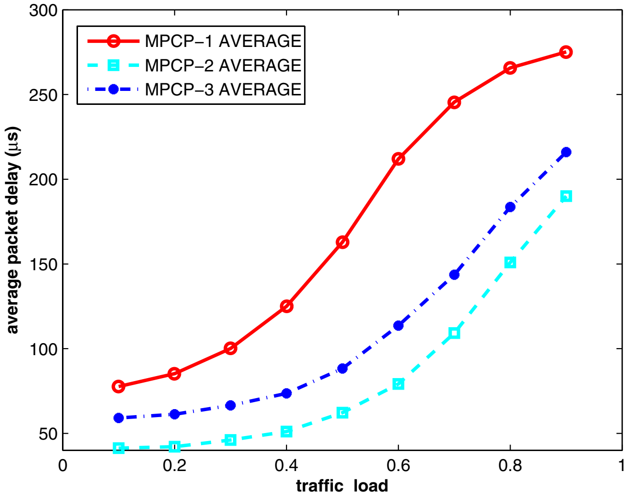

Fig. 11. Packet delay versus traffic load, $N = 128$, $W = 16$, hot-spot traffic.

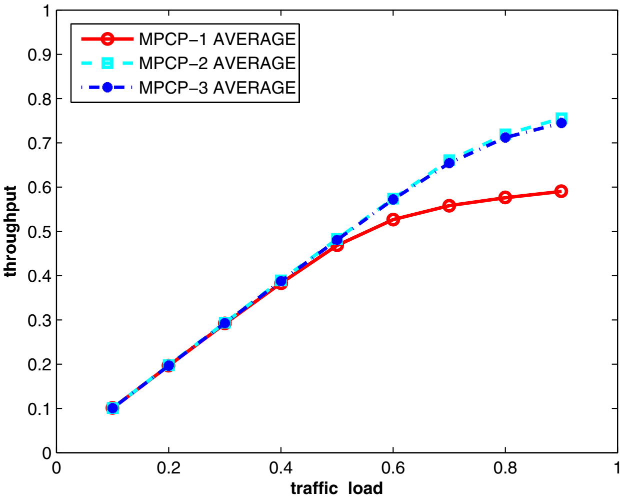

Fig. 13. Throughput versus traffic load, $N = 64$, $W = 8$, hot-spot traffic.

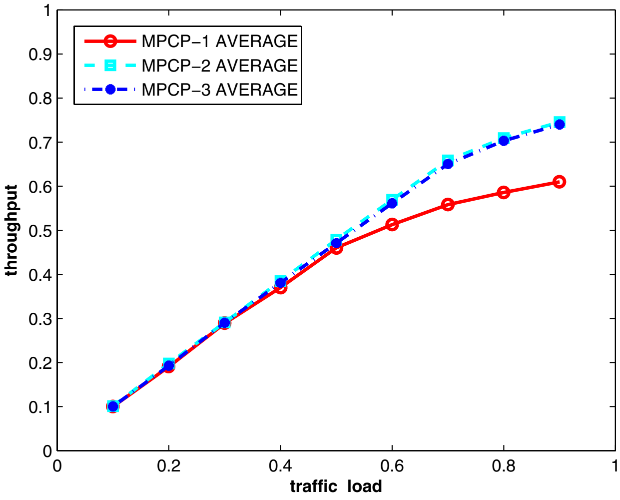

Fig. 14. Throughput versus traffic load, $N = 128$, $W = 16$, hot-spot traffic.

order. Also, the throughput of MPCP-2 and MPCP-3 increases almost linearly even at high load values.

Overall, the results of our simulation study indicate that a relatively small modification to the MPCP protocol allows the OLT to mask the idle time between transmission rounds by performing look-ahead scheduling of ONU requests for bandwidth. This look-ahead feature is applicable to both single-channel and multichannel EPONs. The look-ahead operation requires no changes to the ONUs and can be implemented via a software update to the OLT. This simple modification to the MPCP protocol can be very effective in (1) lowering the average packet delay and (2) allowing the upstream channels to operate at high loads without a significant decrease in their traffic carrying capacity.

V. Concluding Remarks

We have presented MPCP-ℓ, a simple yet effective extension to the MPCP protocol for WDM EPONs that allows for look-ahead scheduling of the upstream channels. Our simulation results indicate that the proposed look-ahead operation increases bandwidth utilization and improves average packet delay. Although this work focused on demonstrating the benefits of the look-ahead operation using a simple bandwidth allocation algorithm, the new look-ahead feature makes it possible to design new, sophisticated DBA schemes that can take advantage of the additional information at the OLT to support advanced QoS capabilities; this is an area of ongoing research in our group.

Acknowledgments

This work was supported in part by the National Science Foundation under Grant CNS-1113191, and in part by the Deanship of Scientific Research (DSR), King Abdulaziz University, under Grant No. 2-611-1434-HiCi.

References

[1] M. McGarry, M. Reisslein, and M. Maier, “WDM Ethernet passive optical networks,” *IEEE Commun. Mag.*, vol. 44, no. 2, pp. 15–22, 2006.

[2] G. Kramer and G. Pesavento, “Ethernet passive optical network (EPON): Building a next-generation optical access network,” *IEEE Commun. Mag.*, vol. 40, pp. 66–73, Feb. 2002.

[3] G. Pesavento, “Ethernet passive optical network (EPON),” *Opt. Netw. Mag.*, vol. 4, no. 1, pp. 107–113, 2003.

[4] D. Xue, Y. Qin, and C. Siew, “Deterministic QoS provisioning with network calculus based admission control in WDM EPON networks,” in *Proc. of IEEE ICC*, June 2009.

[5] V. Sivaraman and G. N. Rouskas, “A reservation protocol for broadcast WDM networks and stability analysis,” *Comput. Netw.*, vol. 32, pp. 211–227, Feb. 2000.

[6] G. Keiser and J. Wiley, *FTTX Concepts and Applications*. Wiley, 2006.

[7] A. Banerjee, Y. Park, F. Clarke, H. Song, S. Yang, G. Kramer, K. Kim, and B. Mukherjee, “Wavelength-division-multiplexed passive optical network (WDM-PON) technologies for broadband access: A review [Invited],” *J. Opt. Netw.*, vol. 4, pp. 737–758, Nov. 2005.

[8] M. Maier, M. Herzog, and M. Reisslein, “STARGATE: The next evolutionary step toward unleashing the potential of WDM EPONs,” *IEEE Commun. Mag.*, vol. 45, pp. 50–56, May 2007.

[9] M. McGarry, M. Reisslein, and M. Maier, “Ethernet passive optical network architectures and dynamic bandwidth allocation algorithms,” *IEEE Commun. Surv. Tutorials*, vol. 10, no. 3, pp. 46–60, 2008.

[10] B. Mukherjee, *Optical WDM Networks*. New York: Springer-Verlag, 2006.

[11] G. Kramer, B. Mukherjee, and G. Pesavento, “IPACT: A dynamic protocol for an Ethernet PON (EPON),” *IEEE Commun. Mag.*, vol. 40, pp. 74–80, Feb. 2002.

[12] G. Kramer, B. Mukherjee, and G. Pesavento, “Interleaved polling with adaptive cycle time (IPACT): A dynamic bandwidth distribution scheme in an optical access network,” *Photon. Netw. Commun.*, vol. 4, no. 1, pp. 89–107, 2002.

[13] B. Lannoo, L. Verslegers, D. Colle, M. Pickavet, M. Gagnaire, and P. Demeester, “Analytical model for the IPACT dynamic bandwidth allocation algorithm for EPONs,” *J. Opt. Netw.*, vol. 6, no. 6, pp. 677–688, 2007.

[14] M. T. Ngo, A. Gravey, and D. Bhadauria, “A mean value analysis approach for evaluating the performance of EPON with gated IPACT,” in *Proc. of ONDM*, 2008.

[15] Y. Zhu and M. Ma, “IPACT with grant estimation (IPACT-GE) scheme for Ethernet passive optical networks,” *J. Lightwave Technol.*, vol. 26, no. 14, pp. 2055–2063, 2008.

[16] H. Byun, J. Nho, and J. Lim, “Dynamic bandwidth allocation algorithm in Ethernet passive optical networks,” *Electron. Lett.*, vol. 39, no. 13, pp. 1001–1002, 2003.

[17] A. Gumaste and I. Chlamtac, “A protocol to implement Ethernet over PON,” in *Proc. of Int. Conf. on Communication*, vol. 2, 2003, pp. 1345–1349.

[18] K. Kwong, D. Harle, and I. Andonovic, “Dynamic bandwidth allocation algorithm for differentiated services over WDM EPONs,” in *Proc. of Int. Conf. on Communications Systems*, 2004, pp. 116–120.

[19] F. Clarke, S. Sarkar, and B. Mukherjee, "Simultaneous and interleaved polling: An upstream protocol for WDM-PON," in *Optical Fiber Communication Conf.*, 2006.

[20] H. Lee, H. Choo, T. Lee, and M. Chung, "Group-synchronized polling algorithm in WDM-EPON," *IEEE Commun. Lett.*, vol. 13, no. 3, pp. 199–201, 2009.

[21] M. McGarry, M. Reisslein, C. Colbourn, M. Maier, F. Aurzada, and M. Scheutzow, "Just-in-time scheduling for multichannel EPONs," *J. Lightwave Technol.*, vol. 26, no. 10, pp. 1204–1216, 2008.

[22] K. Kanonakis and I. Tomkos, "Improving the efficiency of online upstream scheduling and wavelength assignment in hybrid WDM/TDMA EPON networks," *IEEE J. Sel. Areas Commun.*, vol. 28, no. 6, pp. 838–848, 2010.

[23] T. Das, A. Gumaste, A. Lodha, A. Mathew, and N. Ghani, "Generalized framework and analysis for bandwidth scheduling in GPONs and NGPONs—the K-out-of-N approach," *J. Lightwave Technol.*, vol. 29, pp. 2875–2892, Oct. 2011.

[24] Y. Luo and N. Ansari, "Bandwidth allocation for multi-service access on EPONs," *IEEE Commun. Mag.*, vol. 43, no. 2, pp. 516–521, 2005.

[25] C. Assi, Y. Ye, S. Dixit, and M. Ali, "Dynamic bandwidth allocation for quality-of-service over Ethernet PONs," *IEEE J. Sel. Areas. Commun.*, vol. 21, no. 9, pp. 1467–1477, 2003.

[26] J. Xie, S. Jiang, and Y. Jiang, "A dynamic bandwidth allocation scheme for differentiated services in EPONs," *IEEE Commun. Mag.*, vol. 42, no. 8, pp. S32–S39, 2004.

[27] R. Salles and J. Barria, "Fair and efficient dynamic bandwidth allocation for multi-application networks," *Comput. Netw.*, vol. 49, no. 6, pp. 856–877, 2005.

[28] J. Zheng, "Efficient bandwidth allocation algorithm for Ethernet passive optical networks," *IEE Proc. Commun.*, vol. 153, no. 3, pp. 464–468, 2006.

[29] A. Dhaini, C. Assi, M. Maier, and A. Shami, "Dynamic wavelength and bandwidth allocation in hybrid TDM/WDM EPON networks," *J. Lightwave Technol.*, vol. 25, no. 1, pp. 277–286, 2007.

[30] C. Wang, W. Wei, W. Zhang, H. Jiang, C. Qiao, and T. Wang, "Optimal wavelength scheduling for hybrid WDM/TDM passive optical networks," *J. Opt. Commun. Netw.*, vol. 3, no. 6, pp. 522–532, 2011.

[31] R. Sinha, C. Papadopoulos, and J. Heidemann, "Internet packet size distributions: Some observations," Oct. 2005 [Online]. Available: ftp://ftp.isi.edu/isi-pubs/tr-643.pdf.

[32] T. Gonzalez, O. Ibarra, and S. Sahni, "Bounds for LPT schedules on uniform processors," *SIAM J. Comput.*, vol. 6, no. 1, pp. 155–166, 1977.

[33] I. Baldine and G. N. Rouskas, "Traffic adaptive WDM networks: A study of reconfiguration issues," *J. Lightwave Technol.*, vol. 19, pp. 433–455, Apr. 2001.

[34] I. Baldine and G. N. Rouskas, "Reconfiguration and dynamic load balancing in broadcast WDM networks," *Photon. Netw. Commun.*, vol. 1, pp. 49–64, June 1999.

[35] X. Liu and G. N. Rouskas, "MPCP-ℓ: Look-ahead enhanced MPCP for EPON," in *Proc. of IEEE Int. Conf. on Communications*, June 2013.

Xiaomin Liu is a Ph.D. candidate in the School of Electronic and Information Engineering, Beihang University, Beijing, China. She received her B.S. degree from the School of Electronic and Information Engineering, Beihang University, in 2008. Her main research areas are in computer networks, including high-speed communication networks, optical networks, and network analysis and performance evaluation.

George N. Rouskas is a professor of computer science at North Carolina State University and an IEEE fellow. He received his diploma in computer engineering from the National Technical University of Athens (NTUA), Athens, Greece, and his M.S. and Ph.D. degrees in computer science from the College of Computing, Georgia Institute of Technology, Atlanta, GA. His research interests include network architectures and protocols, optical networks, network design and optimization, and performance evaluation. He is coeditor of the book *Next-Generation Internet Architectures and Protocols* (Cambridge University Press, 2011), author of the book *Internet Tiered Services* (Springer, 2009), and coeditor of the book *Traffic Grooming for Optical Networks* (Springer 2008). He is founding coeditor-in-chief of the journal *Optical Switching and Networking*, and he has served on the editorial boards of *IEEE/ACM Transactions on Networking*, IEEE/OSA *Journal of Optical Communications and Networking*, *Computer Networks*, and *Optical Networks*. He is the general chair for ICCCN 2013, and he has served as TPC or general chair for several conferences, including ICCCN 2011, the IEEE GLOBECOM 2010 ONS Symposium, BROADNETS 2007, IEEE LANMAN 2004 and 2005, and IFIP NETWORKING 2004. He is the recipient of a 1997 NSF CAREER Award, the 2004 ALCOA Foundation Engineering Research Achievement Award, and the 2003 NCSU Alumni Outstanding Research Award, and he was inducted into the NCSU Academy of Outstanding Teachers in 2004. He served as the secretary of the IEEE Optical Networking Technical Committee (ONTC) in 2012–2013, and he served as a distinguished lecturer for the IEEE Communications Society in 2010–2011.

Feng He is an instructor in the School of Electronics and Information Engineering at Beihang University, China. He received his Ph.D. degree in communication and information systems from Beihang University in 2009 and his bachelor's degree in electronics and information engineering in 2003. His research interests cover avionics system architecture, scheduling algorithms, and real-time networks.

Huagang Xiong is a full professor in the School of Electronics and Information Engineering at Beihang University, China. He is the vice president of the school and the deputy director of the avionics central laboratory. His research interests include communication network theory and technology practice, avionics information integration, ultra-wideband communication, and electronic device smart testing. He is a member of the Avionics Standardization Committee, a director of the Electronic Circuit Council of Beijing, and a chair fellow in the Chinese Institute of Electronics, Aviation Institute, and Communication Institute. He is also an editor-in-chief of the *Journal of Space Science* and *Electronics Optics and Control*.